ECE 509: Convex Optimization **Rutgers: Spring 2024**

Homework #6

March 21, 2024

Name: Ravi Raghavan *Extension: No*

Problem 1. *Pre-computation for line searches.* For each of the following functions, explain how the computational cost of a line search can be reduced by a pre-computation. Give the cost of the pre-computation, and the cost of evaluating $g(t) = f(x + t\Delta x)$ and $g'(t)$ with and without the pre-computation.

(a) $f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x)$

Solution. In Exact Line Search, our goal is to minimize $\tilde{f}(t) = f(x + t\Delta x)$

$\tilde{f}(t) = -\sum_{i=1}^{m} \log(b_i - a_i^T(x + t\Delta x))$

$\tilde{f}(t) = -\sum_{i=1}^{m} \log(b_i - a_i^T x - a_i^T t\Delta x))$

$\tilde{f}'(t) = -\sum_{i=1}^{m} \frac{-a_i^T \Delta x}{\log(b_i - a_i^T(x + t\Delta x)))}$

$\tilde{f}'(t) = -\sum_{i=1}^{m} \frac{-a_i^T \Delta x}{\log(b_i - a_i^T x - a_i^T t\Delta x))}$

Let n be the size of each a_i and x. Let's say we are testing out k values of t.

Precomputation:

Computing $b_i - a_i^T x$ for $i \in [1, m]$ takes $O(n)$ operations per i.
Computing $a_i^T \Delta x$ for $i \in [1, m]$ takes $O(n)$ operations per i

Hence, there are a total of $m * O(n) = O(mn)$ operations in pre-computation

Cost without Precomputation($\tilde{f}(t)$):
Computing $x + t\Delta x$ takes a total of $O(n)$ operations.

Computing $a_i^T(x + t\Delta x)$ takes a total of $O(n)$ operations.

$b_i - (a_i^T(x + t\Delta x)))$ will take $O(n)$ operations.

$\log(b_i - a_i^T(x + t\Delta x))$ takes $O(n)$

$\sum_{i=1}^{m} \log(b_i - a_i^T(x + t\Delta x))$ takes $O(mn)$ operations.

$-\sum_{i=1}^{m} \log(b_i - a_i^T(x + t\Delta x))$ takes $O(mn)$ operations.

Since we are testing out k values of t, the overall Line Search Algorithm will be $O(mnk)$

Cost without Precomputation($\tilde{f}'(t)$):
Computing $x + t\Delta x$ takes a total of $O(n)$ operations.

Computing $a_i^T(x + t\Delta x)$ takes a total of $O(n)$ operations.

$b_i - (a_i^T(x + t\Delta x)))$ will take $O(n)$ operations.

$\log(b_i - a_i^T(x + t\Delta x))$ takes $O(n)$

Computing $-a_i^T \Delta x$ takes $O(n)$ operations

Calculating $\frac{-a_i^T \Delta x}{\log{(b_i - a_i^T(x + t\Delta x)))}}$ takes $O(n)$

$\sum_{i=1}^{m} \frac{-a_i^T \Delta x}{\log{(b_i - a_i^T(x + t\Delta x)))}}$ takes $O(mn)$ operations.

$-\sum_{i=1}^{m} \frac{-a_i^T \Delta x}{\log{(b_i - a_i^T(x + t\Delta x)))}}$ takes $O(mn)$ operations.

Since we are testing out k values of t, the overall Line Search Algorithm will be $O(mnk)$

Cost with Precomputation($\tilde{f}(t)$):

Computing $\log{(b_i - a_i^T x - a_i^T t\Delta x))}$ takes $O(1)$ operations.

$\sum_{i=1}^{m} \log{(b_i - a_i^T x - a_i^T t\Delta x))}$ takes $O(m)$ operations.

$-\sum_{i=1}^{m} \log{(b_i - a_i^T x - a_i^T t\Delta x))}$ takes $O(m)$ operations.

Since we are testing out k values of t, the Line Search Algorithm will be $O(mk)$
Including pre-computation, it takes $O(mk) + O(mn)$

Cost with Precomputation($\tilde{f}'(t)$):

Computing $\frac{-a_i^T \Delta x}{\log{(b_i - a_i^T x - a_i^T t\Delta x))}}$ takes $O(1)$ operations.

$\sum_{i=1}^{m} \frac{-a_i^T \Delta x}{\log{(b_i - a_i^T x - a_i^T t\Delta x))}}$ takes $O(m)$ operations.

$-\sum_{i=1}^{m} \frac{-a_i^T \Delta x}{\log{(b_i - a_i^T x - a_i^T t\Delta x))}}$ takes $O(m)$ operations.

Since we are testing out k values of t, the Line Search Algorithm will be $O(mk)$
Including pre-computation, it takes $O(mk) + O(mn)$

(b) $f(x) = \log(\sum_{i=1}^{m} \exp(a_i^T x + b_i))$

Solution. In Exact Line Search, our goal is to minimize $\tilde{f}(t) = f(x + t\Delta x)$

$\tilde{f}(t) = \log(\sum_{i=1}^{m} \exp(a_i^T(x + t\Delta x) + b_i))$

$\tilde{f}(t) = \log(\sum_{i=1}^{m} \exp(a_i^T x + a_i^T t\Delta x + b_i))$

$\tilde{f}(t) = \log(\sum_{i=1}^{m} \exp(a_i^T x + b_i + a_i^T t\Delta x))$

$\tilde{f}'(t) = \log(\sum_{i=1}^{m} \exp(a_i^T(x + t\Delta x) + b_i)(a_i^T \Delta x))$

$\tilde{f}'(t) = \log(\sum_{i=1}^{m} \exp(a_i^T x + b_i + a_i^T t\Delta x)(a_i^T \Delta x))$

Let n be the size of each a_i and x. Let's say we are testing out k values of t.

Precomputation:

Computing $a_i^T x + b_i$ for $i \in [1, m]$ takes $O(n)$ operations per i.
Computing $a_i^T \Delta x$ for $i \in [1, m]$ takes $O(n)$ operations per i

Hence, there are a total of $m * O(n) = O(mn)$ operations in pre-computation

Cost without Precomputation($\tilde{f}(t)$):

Computing $x + t\Delta x$ takes a total of $O(n)$ operations.

Computing $a_i^T(x + t\Delta x)$ takes a total of $O(n)$ operations.

$a_i^T(x + t\Delta x) + b_i$ takes $O(n)$ operations.

$\exp(a_i^T(x + t\Delta x) + b_i)$ takes $O(n)$

$\sum_{i=1}^m \exp(a_i^T(x + t\Delta x) + b_i)$ takes $O(mn)$ operations.

$\log(\sum_{i=1}^m \exp(a_i^T x + b_i + a_i^T t\Delta x))$ takes $O(mn)$ operations

Since we are testing out k values of t, the overall Line Search Algorithm takes $O(mnk)$

Cost without Precomputation($\tilde{f}'(t)$):
Computing $x + t\Delta x$ takes a total of $O(n)$ operations.

Computing $a_i^T(x + t\Delta x)$ takes a total of $O(n)$ operations.

$a_i^T(x + t\Delta x) + b_i$ takes $O(n)$ operations.

$\exp(a_i^T(x + t\Delta x) + b_i)$ takes $O(n)$

$\exp(a_i^T(x + t\Delta x) + b_i)(a_i^T\Delta x)$ takes $O(n)$

$\sum_{i=1}^m \exp(a_i^T(x + t\Delta x) + b_i)$ takes $O(mn)$ operations.

$\log(\sum_{i=1}^m \exp(a_i^T(x + t\Delta x) + b_i)(a_i^T\Delta x))$ takes $O(mn)$ operations

Since we are testing out k values of t, the overall Line Search Algorithm takes $O(mnk)$

Cost with Precomputation($\tilde{f}(t)$):
Computing $\exp(a_i^T x + b_i + a_i^T t\Delta x)$ takes $O(1)$ operations.

$\sum_{i=1}^m \exp(a_i^T x + b_i + a_i^T t\Delta x)$ takes $O(m)$ operations.

$\log(\sum_{i=1}^m \exp(a_i^T x + b_i + a_i^T t\Delta x))$ takes $O(m)$ operations

Since we are testing out k values of t, the Line Search Algorithm takes $O(mk)$
Including pre-computation, it takes $O(mk) + O(mn)$

Cost with Precomputation($\tilde{f}'(t)$):
Computing $\exp(a_i^T x + b_i + a_i^T t\Delta x)(a_i^T\Delta x)$ takes $O(1)$ operations.

$(\sum_{i=1}^m \exp(a_i^T x + b_i + a_i^T t\Delta x)(a_i^T\Delta x))$ takes $O(m)$ operations.

$\log(\sum_{i=1}^m \exp(a_i^T x + b_i + a_i^T t\Delta x)(a_i^T\Delta x))$ takes $O(m)$ operations

Since we are testing out k values of t, the Line Search Algorithm takes $O(mk)$
Including pre-computation, it takes $O(mk) + O(mn)$

Problem 2. *True or False.*

(a) False

(b) True

(c) False

(d) True

(e) False

(f) True

(g) False